

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
26 September 2002 (26.09.2002)

PCT

(10) International Publication Number
WO 02/076042 A1

(51) International Patent Classification⁷: **H04L 12/56**

(21) International Application Number: **PCT/IB02/00822**

(22) International Filing Date: 15 March 2002 (15.03.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
01810264.0 19 March 2001 (19.03.2001) EP

(71) Applicant (for all designated States except US): **INTERNATIONAL BUSINESS MACHINES CORPORATION [US/US]**; New Orchard Road, Armonk, NJ 10504 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **LUIJTEN, Ronald, P.** [NL/CH]; Saentisstrasse 23, CH-8810 Horgen (CH). **HERKERSDORF, Andreas** [DE/CH]; Ruetihaldenstrasse 10, CH-8135 Langnau a. A. (CH).

(74) Agent: **KLETT, Peter, Michael**; International Business Machines Corporation, Saeumerstrasse 4/Postfach, CH-8803 Rueschlikon (CH).

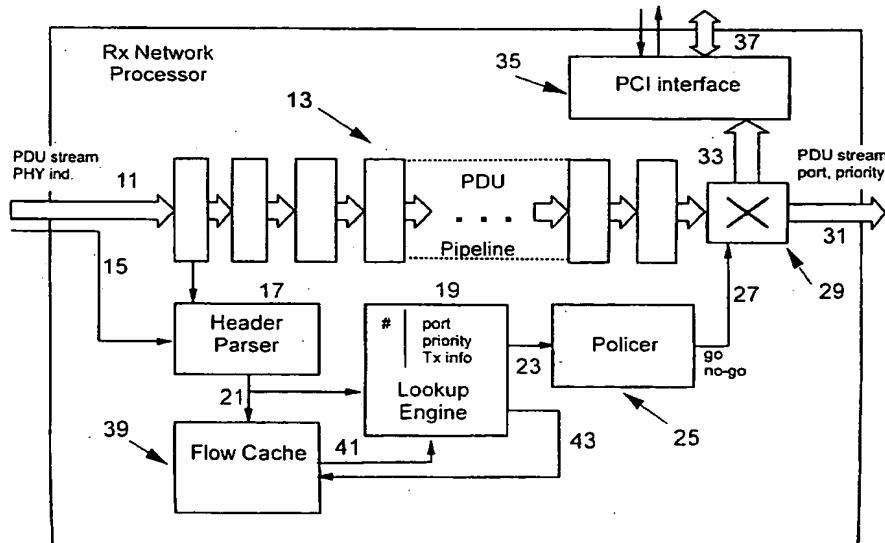
(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— with international search report

[Continued on next page]

(54) Title: CACHE ENTRY SELECTION METHOD AND APPARATUS



WO 02/076042 A1

(57) Abstract: A method for selectively inserting cache entries into a cache memory is proposed in which incoming data packets are directed to output links according to address information. The method comprises the following steps: a) an evaluation step for evaluating for each incoming data packet classification information which is relevant to the type of traffic flow or to the traffic priority to which the data packet is associated; b) a selection step for selecting based on the result of the evaluation step whether for the data packet the cache entry is to be inserted into the cache memory; c) an entry step for inserting as the cache entry into the cache memory, in the case the result of the selection step is that the cache entry is to be inserted, for the data packet the address information and associated output link information.



— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

- 1 -

CACHE ENTRY SELECTION METHOD AND APPARATUS

Field of Invention

The present invention is concerned with the routing of data packets in a communication network, according to addresses (or directional information) contained in the header of each 5 data packet. In particular, it is concerned with handling of large quantities of data packets in routing nodes in which caches are used for storing previously used addresses together with associated routing information, to avoid the lookup operation in a routing table for each occurrence of frequently used data packet addresses. As a data packet may be seen any information unit, e.g. as processed by CPUs.

10

Background

In modern communication systems based on data packet transmission, such as the Internet, millions of data packets per second are forwarded in each switching node from the input to one of several output links, depending on the address information contained in the data packet header. The association between data packet address and output link (or route) is stored in a 15 routing table of the switching node, and for each data packet a lookup or search operation is made in this table for obtaining the routing information, also referred to as output link identification. This lookup may cause delays if resources are limited. As many data packets - such as periodic data packets of voice streams, or bursts of data packets for file transfers - contain the same address, one has used cache memories for storing each addresses when it first 20 occurred and the respective associated output or routing information once it has been found by a lookup operation. This avoids many recurring lookup searches for the same address, but requires large sizes of cache memories. However, due to the finite size of any cache memory, storage locations have to be cleared for new cache entries once the cache is filled up. Often, an LRU (least recently used) algorithm is used for this task, thus eliminating cache entries which 25 obviously were not used for some time.

However, when the cache memory management employs an LRU algorithm, and when the communication system has to transfer a great number of data packets for traffic of periodic nature such as voice or video transmissions, many intervening data packets will occur between consecutive data packets of the same periodic transmission, so that most cache entries would

- 2 -

have been eliminated before they can be used again. This is of course very ineffective. Furthermore, some traffic categories should be handled with higher priority than others. Thus, some kind of optimization or improvement is required for the cache entry procedure to avoid the handling effort and the waste of cache space for cache entries which will be used rarely or
5 only once anyway.

The usage of cache memories in routers of IP data packet forwarding systems has been discussed in the following publication: Peter Newman et al. : "IP Switching and Gigabit Routers", IEEE Communications Magazine, Vol. 35, No. 1 (January 1997), pp. 64-69. The result of this investigation was that caching for high-speed links at 1Gbit/s and above will not
10 be useful due to the lack of locality in addresses. Another approach discussed in this paper which deals with a particular (Ipsilon) IP switching system is conceptually different in that it suggests to forward connection-less IP traffic over a connection-oriented ATM traffic infrastructure.

Classification of data packets in IP switching nodes in response to the contents of their headers
15 was discussed in T. V. Lakshman, D. Stiliadis: "High-speed Policy-based Packet Forwarding Using Efficient Multi-dimensional Range Matching", Proc. ACM SIGCOMM'98, Comp. Commun. Rev. Vol. 28, No. 4, Oct. 1998, pp. 203-214 and also in P. Gupta, N. McKeown: "Packet Classification on Multiple Fields", ACM SIGCOMM'99, Comp. Commun. Rev. Vol. 29, No. 4, Ovt. 1999, pp. 147-160.

20 In none of these publications, procedures were suggested for connection-less data packet transmission systems to select, in response to the evaluation of data packet headers, certain flows or traffic categories for entry into a cache memory, to avoid later lookup operations for data packets of the same flow or traffic category, but to avoid on the other hand useless cache entries into the cache.

25

Objects of the Invention

It is an object of the invention to improve the usage of the cache memory.

This is of particular advantage in data-packet-handling devices of switching nodes or routers. In particular, those cache entries which will be used with higher probability than others, or cache

- 3 -

entries for traffic types which suffer more from the time delay due to routing table lookup than others, should be privileged or preferred in favor of other cache entries for e. g. periodic traffic.

Summary of the Invention

A method for inserting cache entries into a cache memory is proposed in which incoming data
5 packets are directed to output links according to address information. The method comprises the following steps:

- a) an evaluation step for evaluating for each incoming data packet classification information which is relevant to the type of traffic flow or to the traffic priority to which the data packet is associated;
- 10 b) a selection step for selecting based on the result of the evaluation step whether for the data packet the cache entry is to be inserted into the cache memory;
- c) an entry step for inserting as the cache entry into the cache memory, in the case the result of the selection step is that the cache entry is to be inserted, for the data packet the address information and associated output link information.
- 15 The method allows to keep cache entries for such data packet flows where it is foreseeable that additional data packets of the same data packet flow will follow within a relatively short time period, measured in data packet arrival rate, instead of data packet entries for traffic flows that are more or less evenly distributed or spaced in time e.g. with lots of data packets between two consecutive data packets of the same flow. However, periodic data packet flows which are
20 specially identified as e. g. having a high priority, can nevertheless be selected for caching with this mechanism.

Thus, employment of this invention has following advantages: The size of the cache can be smaller than with ordinary cache entry procedures, because the cache entries can be restricted to those candidates, i.e. traffic flows which have a high probability of future cache hits. Traffic
25 with periodic or even occasional transmissions which would need a normal lookup procedure anyway, because their cache entries would be eliminated in the meantime between two occurring data packets of the same flow, will not burden the cache memory space and the units handling cache entries.

The invention can be particularly useful in communication systems based on data packet transmission, where data packets are forwarded in switching nodes or routers from an input to one of several output links, depending on address information contained in the data packet header. The association between data packet address and output link or route is stored in a routing table of the switching node, and for each data packet a lookup or search operation is made in this table for obtaining the routing information, also referred to as output link information. Typically data packets that together form a data packet flow are selected for inserting a cache entry therefor into the cache memory.

The concept is not limited to data packet transmission systems but can be used in, or be adapted to, other systems in which reoccurring accesses to search mechanisms is performed, such as systems in which databases are accessed by different users with different frequency or urgency. In computer systems there exists also a data cache where this invention might be advantageously applied, e.g. in the context of a branch-through table, whose data should not be inserted as cache entry in the cache memory because a soon further access to the same data is improbable, whereas an often used variable for a program would be a case for which a cache entry is recommended.

Listing of Drawings

Embodiments of the invention will be described with reference to the accompanying drawings.

In the drawings:

20 Fig. 1 is a block diagram of an Rx network processor in a switching node or router, comprising a lookup engine for performing lookup operations in a routing table and an IP switching unit, as well as a data packet header parser and a cache memory, in which the header parser and the lookup engine cooperate for selecting addresses and output link information of certain flows for cache entry;

25 Fig. 2 shows a data packet header for schematically representing the operations of the header parser for making cache entry decisions when a priority is explicitly indicated;

- 5 -

Fig. 3 shows a data packet header for schematically representing the operations of the lookup engine 19 containing the routing table, in connection with the header parser, for making cache entry decisions for cases where a priority is implicitly encoded,

5 Fig. 4 illustrates a data packet header of bulk data packet transmissions in TCP data flows, where the start and end of a data packet group are deducible from SYN and FIN indications, for making cache entry decisions for such flows.

Detailed Description of the Embodiments

Network Processor

Fig. 1 schematically shows a network processor in a node of a communication system, in which the headers of incoming data packets, also referred to as Protocol Data Units or PDUs, are analyzed so that each data packet can be directed to one of several output ports, or suppressed if the present policy requires this.

The stream of data packets is received on a bus 11 and each data packet is entered into a PDU pipeline 13 to reserve time for the header evaluation step. Each data packet header is entered via an input 15 into a header parser 17 which extracts relevant information, also referred to as classification information, such as destination address, priority, class of service, etc. from the header and transfers it to a lookup engine 19 via a connection 21. The lookup engine 19 contains a lookup table, also referred to as routing table or rule table, in which for each destination address etc. the respective output port (target port) of the node and other handling information is stored. A lookup step using the routing table is performed. The lookup result is transferred through lines 23 via a policer 25 and lines 27 to a switching unit 29 which receives the data packets from the pipeline 13 and directs each data packet, in response to the routing information received from the lookup engine 19 and the policer 25, to the correct target port over connections 31. Depending on the policies stored in the policer 25, it can either send the output link information obtained from lookup engine 19 to the switching unit 29, or a no-go signal so that the further transfer of a data packet (PDU) can be suppressed if the policy requires that.

The data packets can also be transferred through lines 33 to a PCI interface 35 for further distribution over connections 37 to a control processor (not shown) for special handling.

To save the time that is used for the lookup operations, a cache memory 39 is provided into which output information such as target port ID, etc. that was found in a previous lookup operation is stored together with the input (destination address). When a destination address or other relevant input information appears on the lines 21, it is also transferred to the cache memory 39, and if it is found in the cache memory 39, the respective output information is transferred through lines 41 to the lookup engine 19 which then can stop the lookup operation

- 7 -

and transfer the output information on lines 23 to the policer 25 and the switching unit 29. If the destination address was not yet stored in the cache memory 39, the corresponding output information found in a search operation is received from the lookup engine 19 over lines 43, so that both can be stored in the cache memory 39.

5 As the cache memory 39 is of limited size, it would be filled up after some time and old cache entries should be canceled to allow storage of new cache entries. The well-known least-recently used (LRU) algorithm can be used for selecting cache entries to be canceled. However, as was mentioned in the introduction, the LRU algorithm may be an inadequate procedure: For periodic transmissions, there will be many intervening data packets between two sequential
10 data packets of the same flow, so that many cache entries of the cache memory 39 will be canceled before they can be used again. Therefore, another solution for optimizing the utilization of cache memory 39 space is introduced.

Selection Procedure for Caching

Cache entries are made for data packets selected in a selection step, e. g. those of higher
15 priorities or having other criteria such as a start of flow identifier which hints on subsequently expected data packets with identical address (e. g. TCP - SYN). A group of such data packets that belong together in that sense are referred to as a data packet flow. The selection step could be made, as illustrated in Fig. 2 for a first embodiment, by the header parser 17 which sends a respective signal to the cache memory 39 and to the lookup engine 19 over lines 21. Another
20 solution as illustrated in Fig. 3 for a second embodiment is to make the selection depending on the result of a lookup operation in the lookup engine 19 which then sends a respective selection signal over lines 43 to the cache memory 39. This second scheme is used when, e. g., the data packet priority cannot be explicitly deduced from the data packet header but is determined during a data packet classification process. A third embodiment for a cache entry decision
25 comprising evaluation and selection is illustrated in Fig. 4: The header parser 17 evaluates in an evaluation step data packet headers to detect the start or end of a TCP (bulk) transmission. If such a start or end is detected, this result is used for determining whether for a data packet a cache entry shall be inserted or not.

Operation with the selection mechanism is as follows:

- 8 -

The header of each incoming data packet is entered into the header parser 17. The relevant information, e.g. destination address, is transferred to the cache memory 39 and the lookup engine 19 which both start their operation in parallel to find the required output information (target port selection). In a checking step it is checked whether the address information or 5 associated output link information of the data packet is already contained in the cache memory 39.

The header parser 17 and/or the lookup engine 19 also evaluate received information to determine whether the corresponding data packet (PDU) is of a category or data flow that deserves inserting an entry into the cache memory 39. This is the evaluation and selection step.

- 10 (a) If the input and corresponding output information are found during the checking step in the cache memory 39, the operation of the lookup engine 19 is stopped; this result is transferred over lines 41 through the lookup engine 19 and the policer 25 to the switching unit 29. The data packet is forwarded in a forwarding step under use of the output information that has been found.
- 15 (b) If the input information is not found in the cache memory 39, the lookup engine 19 completes its lookup step and, besides transferring the result of this lookup step (target port identification) via the policer 25 to the switching unit 29, also provides it to the cache memory 39 through the lines 43. This cache memory 39 stores in an entry step the input information (destination address) and corresponding output link information (target port ID) in a cache 20 entry if and only if it receives a select-signal from the header parser 17 or from the lookup engine 19 on the lines 21 or 43, respectively.

Details of Different Cache Entry Selection Steps

The selection decision for inserting a cache entry can be based on various criteria, such as address information or communication protocol information about the applied communication 25 protocol. In the following, three examples are given for a selection based either on priority of the received data packets (PDUs), or on an indication of a data packet bulk (group) transmission. In one example, the data packet priority is directly indicated in the data packet header. In another example there is an indirect priority indication which is derived from the

- 9 -

header contents. In the third example, indications of first and last data packets of TCP transmissions are used as selection criterion.

1) Cache entry selection based on explicit priority encoding:

Fig. 2 shows the header of a data packet in which those portions are marked which are (a) evaluated as criterion for making a cache entry decision, and which are (b) used as input value for the lookup operation. These portions represent the classification information which allows to evaluate the data packet and to classify it whether it belongs to the data packets for which a cache entry is inserted or not. The decision criterion is an explicit encoding of the priority of the respective data packet and of all the other data packets which are part of the same data flow.

5 This criterion may be either DiffServ DSCP or VLAN priority bits. The header parser 17 evaluates these bits and depending on the result signals to the lookup engine 19 and to the cache memory 39 whether the search result is to be cached or not.

10

The destination address from the header is extracted by the header parser 17 and transferred to the lookup engine 19 which then performs a longest matching prefix search in its (IPv4) routing table, and provides a respective output value (target port identification) on its outputs on the lines 23, 43. The lookup engine 19 then triggers the entry of the destination address and of the search result (target port and possibly queue priority) into the cache memory 39. However, if the respective destination address had been found in the cache memory 39 already, the lookup operation would have been interrupted, and the output result, i.e. target port etc., will be

15 provided to the lookup engine 19 over the lines 41, and transferred from there through the lines 23 and the policer 25 to the switching unit 29.

20

In this example, the header parser 17 makes the cache entry decision based on the explicit priority indication, and the lookup engine 19 only makes a longest matching prefix search for obtaining the required target port identification (and possibly an assigned priority value).

25 2) Cache entry selection based on implicit priority encoding:

Fig. 3 also shows a data packet (PDU) header in which other fields are marked, i.e. hashed, which are used for evaluation: a five-tuple consisting of (a) protocol indicator, (b) source address, (c) destination address, (d) source port, and (e) destination port. These values are extracted by the header parser 17 and provided to the lookup engine 19. The lookup engine 19

- 10 -

contains the rule table of which an example is shown in the table at the end of present description, by which a classification can be made on the basis of protocol/address five-tuples. The result of the search, i.e. lookup, in the rule table provides two values: (1) a classification (high / low / deny) which determines the priority of the respective data packet. This priority
5 value is provided to the switching unit 29. It also determines whether a cache entry must be made in the cache memory 39, e. g. cache entries to be made only for a high priority. (2) The lookup in the rule table further provides the identification of the target port to be used for the respective data packet.

If a cache entry is to be made into the cache memory 39 for a data packet, this will contain the
10 five-tuple extracted from the header, and the classification result plus the target port identification. As in the first example, if the cache memory 39 already contained the respective five-tuple, the search in the lookup engine 19 is interrupted, and the priority value and destination port ID are transferred from the cache memory 39 to the switching unit 29 through the lookup engine 19 and the policer 25.

15 Thus, in this second example, the header parser 17 only extracts five relevant fields from the data packet header, whereas the lookup engine 19, besides finding the required target port, also makes the cache entry decision.

3) Other decision principles for cache entry selection

Application of the invention is of course not limited to cases as given in the two examples
20 above. Any single header field or combination of header fields which reflect the traffic or flow category of the data packet or its priority can be taken as decision criteria, e. g. MPLS labels, ATM UPI/VCI labels, Ethernet MAC addresses, etc.

Another option for selecting a data packet flow for cache entry, i. e. for entering data packet address and output port information into the cache memory 39, is described below in a third
25 embodiment example with reference to Fig. 4. The respective situation is when a number of data packet arrivals from the same flow can be expected in almost consecutive order or at least with relatively small gaps between individual data packets (locality or burstiness of data packets of the same flow). Such circumstances or conditions can, for example, be deduced from individual data packet headers in the following way, c.f. Fig. 4: TCP transmissions typically

start with a so called SYN message to synchronize the sequence numbers between transmitting and receiving terminal. SYN messages are recognized by a particular control bit (SYN, marked in Fig. 4) set in the TCP header, and are hence easily to parse or detect. The header parser 17 (see Fig. 1) will send a cache entry control signal to the cache memory 39, and the destination port address of the SYN message as well as the output link information, also referred to as address information, provided by the lookup operation for the SYN message will then be entered into the cache memory 39. There is a high probability that subsequent data packets will arrive with the same address information so that output link information can directly be taken from the cache memory 39 for each subsequent data packet of the bulk transmission. This applies hence to burst-type traffic of data packets.

The same scheme can also be applied to explicitly remove respective cache entries from the cache memory 39 and free the corresponding space for new cache entries, as will be explained below. The apparatus for controlling the insertion of cache entries into the cache memory 39, in which incoming data packets are directed to output links according to address information, hence comprises the header parser 17 which here unifies three functionalities: a) an evaluator for evaluating for each incoming data packet classification information which is relevant to the type of flow or to the priority to which the data packet is associated, b) a selector for selecting based on the result of the evaluation step whether for the data packet the cache entry is to be inserted into the cache memory 39, and c) a cache read/write unit for inserting as the cache entry into the cache memory 39, in the case the result of the selection step is that the cache entry is to be inserted, for the data packet the address information and associated output link information. All these functionalities need not be unified but can be performed by separate units also. The apparatus here further comprises an output unit 29 for forwarding the data packet under use of output link information stored in that cache entry, if the address information or associated output link information is found in a cache entry.

Deletion of Entries from the Cache Memory

Even when the described method is used, the cache memory 39 might eventually be filled up, and a cache entry should be deleted from the cache memory 39 to clear a storage position before another cache entry can be inserted. The rule for cache entry deletion can be independent

- 12 -

of the cache entry decision mechanism used, but it could also use the same selection scheme. The LRU rule can be used of course for clearing cache positions.

Another mechanism which is suggested as advantageous for cache position clearing in connection with the cache entry decision mechanism of the third selection example described
5 above (for TCP bulk data packet transmissions) is as follows (cf. also Fig. 4):

The last data packet of one TCP session has a FIN control bit (marked in Fig. 4) set in the TCP header. This FIN message is straightforward to parse or detect in the header parser 17 (Fig. 1), and the header parser 17 will send a control signal for deleting a respective cache entry in the cache memory 39 having the same destination port address as found in the TCP FIN message,
10 because no more data packets with this address are expected to arrive.

Summary of Advantages Gained

As was already shortly mentioned in the introduction, the major advantages that are gained when the described method is employed in a communication system transmitting addressed data packets to various destinations are:

15 For data packet flows which have a high temporal locality of data packets, or which have a particular priority, or which have other selectable criteria, cache entries into a cache memory are inserted. This will speed up the operation of data packet transfers in switching nodes compared to other systems using a normal, i.e. unselected cache entry mechanism, while ensuring that the sizes of cache memories can be smaller as otherwise required for systems
20 using normal cache entry.

- 13 -

Example for a Classification Rule Table

<----- Input = 5-tuple ----->						<-Search	Result-->	
(Each Field: Either a Range or a Single Value)								
No.	SA	DA	SP	DP	PROT	Class	Port	
5	00000000	84c90000	006f	0000	6- 6	deny	-	
	- ffffffff	- 84c9ffff	- 006f	- ffff				
10	00000000	9bb30000	006f	0000	6- 6	deny	-	
	- ffffffff	- 9bb3ffff	- 006f	- ffff				
	00000000	84c90000	006f	0000	11-11	deny	-	
	- ffffffff	- 84c9ffff	- 006f	- ffff				
15	00000000	9bb30000	006f	0000	11-11	deny	-	
	- ffffffff	- 9bb3ffff	- 006f	- ffff				
	00000000	84c90000	0801	0000	6- 6	deny	-	
	- ffffffff	- 84c9ffff	- 0801	- ffff				
20	00000000	9bb30000	0801	0000	6- 6	deny	-	
	- ffffffff	- 9bb3ffff	- 0801	- ffff				
	00000000	84c90000	0801	0000	11-11	deny	-	
	- ffffffff	- 9bb3ffff	- 0801	- ffff				
25	00000000	84c90000	0000	0000	1- 1	high	7	
	- ffffffff	- 84c9ffff	- ffff	- ffff				
10	00000000	9bb30000	0000	0000	1- 1	low	4	
	- ffffffff	- 9bb3ffff	- ffff	- ffff				
11	00000000	84c92a0a	0019	0000	6- 6	low	4	
	- ffffffff	- 84c92a0a	- 0019	- ffff				
30	12	00000000	84c92a0a	0035	0000	6- 6	low	7
	- ffffffff	- 84c92a0a	- 0035	- ffff				
13	00000000	84c92a0a	0035	0000	11-11	low	9	
	- ffffffff	- 84c92a0a	- 0035	- ffff				
35	14	00000000	84c92a0a	0400	0000	11-11	low	6
	- ffffffff	- 84c92a0a	- ffff	- ffff				
15	00000000	84c92ab8	0035	0000	6- 6	low	3	
	- ffffffff	- 84c92ab8	- 0035	- ffff				

C L A I M S

1. A method for inserting cache entries into a cache memory (39), in a device in which incoming data packets are directed to output links according to address information, the method comprising the following steps:

- 5 a) an evaluation step for evaluating for each said incoming data packet classification information which is relevant to the type of flow or to the priority to which said data packet is associated,
- b) a selection step for selecting based on the result of said evaluation step whether for said data packet said cache entry is to be inserted into said cache memory (39),
- c) an entry step for inserting as said cache entry into said cache memory (39), in the case the 10 result of said selection step is that said cache entry is to be inserted, for said data packet said address information and associated output link information.

2. Method according to claim 1, further comprising

- a) a checking step for checking for each data packet, whether its address information or associated output link information is already contained in a cache entry,
- 15 b) a forwarding step for forwarding said data packet under use of output link information stored in that cache entry, if the address information or associated output link information is found in a cache entry,

wherein if the address information or associated output link information is not found in a cache entry, the evaluation step and the selection step are performed for said data packet.

- 20 3. Method according to claim 1 or 2, wherein in the case a cache entry is to be made, said output link information is obtained from a lookup step using a routing table.

- 15 -

4. Method according to claim 3, wherein said lookup step is started together with the checking step and is stopped if the address information or associated output link information is found in a cache entry.

5. Method in accordance with one of claims 1 to 4, wherein data packets associated with burst-type traffic are selected for inserting a corresponding cache entry into said cache memory (39).

6. Method in accordance with one of claims 1 to 5, wherein those data packets that have a data packet header field identifying said data packet as a starting or synchronizing data packet of a sequence of related data packets, are selected for inserting a corresponding cache entry into said cache memory (39).

7. Method in accordance with one of claims 1 to 6, wherein the selection for inserting a corresponding cache entry into said cache memory (39) is made on the basis of criteria contained in the header of the data packet.

8. Method in accordance with claim 7, wherein the criteria include address information and communication protocol information.

9. Method in accordance with one of claims 1 to 8, wherein the selection of a data packet for cache entry is made also in the same unit that determines the output link information.

10. Method in accordance with one of claims 1 to 9, further comprising a removal decision step wherein a decision whether to remove a cache entry from the cache memory (39) is made in dependence of the type of flow to which the respective cache entry is associated, and in dependence of the cache filling grade.

11. Apparatus for controlling the insertion of cache entries into a cache memory (39) in a device, in which incoming data packets are directed to output links according to address information, the apparatus comprising:

- 16 -

- a) an evaluator (17) for evaluating for each said incoming data packet classification information which is relevant to the type of flow or to the priority to which said data packet is associated,
- b) a selector (17) for selecting based on the result of said evaluation step whether for said data packet said cache entry is to be inserted into said cache memory (39),

5 c) a cache read/write unit (17) for inserting as said cache entry into said cache memory (39), in the case the result of said selection step is that said cache entry is to be inserted, for said data packet said address information and associated output link information.

12. Apparatus according to claim 11, further comprising an output unit (29) for forwarding said data packet under use of output link information stored in a cache entry, if the address

10 information or associated output link information is found in that cache entry.

13. Apparatus according to claim 11 or 12, further comprising a routing table.

INTERNATIONAL SEARCH REPORT

Int'l Application No
PCT/IB 02/00822

Patent document cited in search report		Publication date		Patent family member(s)		Publication date
US 2001028651	A1	11-10-2001	JP	2001290718 A		19-10-2001
EP 1035687	A	13-09-2000	JP	3199051 B2		13-08-2001
			JP	2000232479 A		22-08-2000
			EP	1035687 A2		13-09-2000

This Page Blank (uspto)

INTERNATIONAL SEARCH REPORT

Inter	Application No
PCT/IB 02/00822	

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 H04L12/56

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>CISCO SYSTEMS: "NetFlow Services and Applications" WHITE PAPER, 'Online! - 1999 pages 1-27, XP002202296 Retrieved from the Internet: <URL: http://www.cisco.com/warp/public/cc/pd/iosw/isoft/neflct/tech/napps_wp.pdf> 'retrieved on 2002-06-12! page 1, column 1, line 19 - line 25 page 2, column 2, line 6 - line 22 page 10, column 2, line 20 - line 35 --- -/-</p>	1-13

 Further documents are listed in the continuation of box C. Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *&* document member of the same patent family

Date of the actual completion of the international search

14 June 2002

Date of mailing of the international search report

12/08/2002

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel (+31-70) 340-2040, Tx. 31 651 epo nl,
 Fax (+31-70) 340-3016

Authorized officer

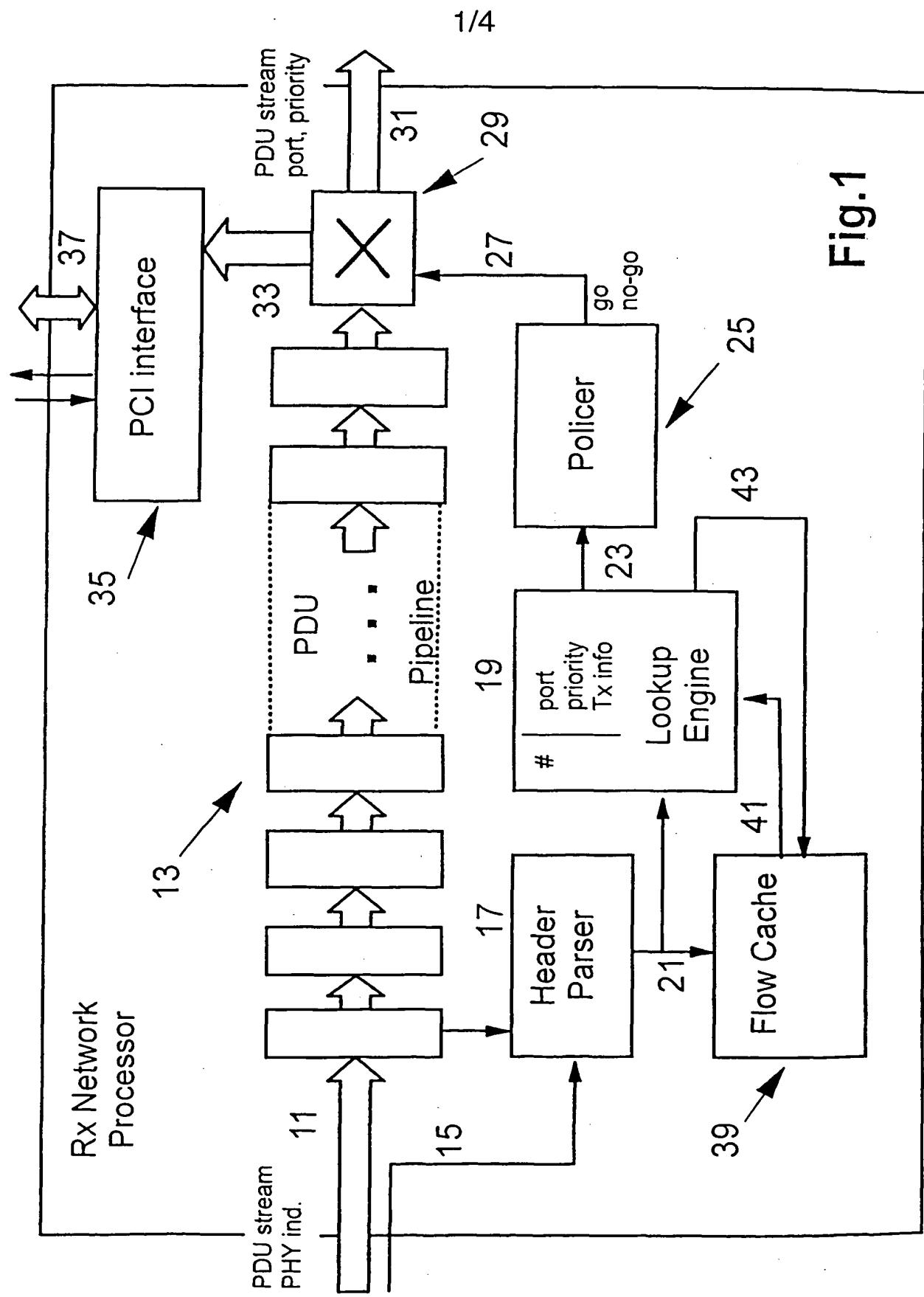
Tous Fajardo, J

INTERNATIONAL SEARCH REPORT

Inter	Application No
PCT/IB 02/00822	

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>YANG B ET AL: "Design, Simulation and Evaluation of TCP Switch" PAPER, 'Online! 4 December 2000 (2000-12-04), pages 1-26, XP002202297 Retrieved from the Internet: <URL:http://www.stanford.edu/class/cs344/staff/group1/final_report.ps> 'retrieved on 2002-06-12! page 5, line 14 -page 6, line 4 ---</p>	1-13
P, X	<p>US 2001/028651 A1 (MURASE TUTOMU) 11 October 2001 (2001-10-11) paragraph '0026! paragraph '0029! paragraph '0033! paragraph '0054! - paragraph '0057! ---</p>	1-13
A	<p>MCCLELLAN R: "Evaluating Expected Network Performance Based on Multilayer Switch Performance Data" TECHNOLOGY WHITE PAPER - MCCLELLAN CONSULTING, 'Online! August 2000 (2000-08), pages 1-13, XP002202298 Retrieved from the Internet: <URL:http://www.mcclellanconsulting.com/whitepapers/perfwpfinal.pdf> 'retrieved on 2002-06-12! page 6, line 38 -page 7, line 11 ---</p>	1-13
A	<p>EP 1 035 687 A (NIPPON ELECTRIC CO) 13 September 2000 (2000-09-13) the whole document -----</p>	1-13



2/4

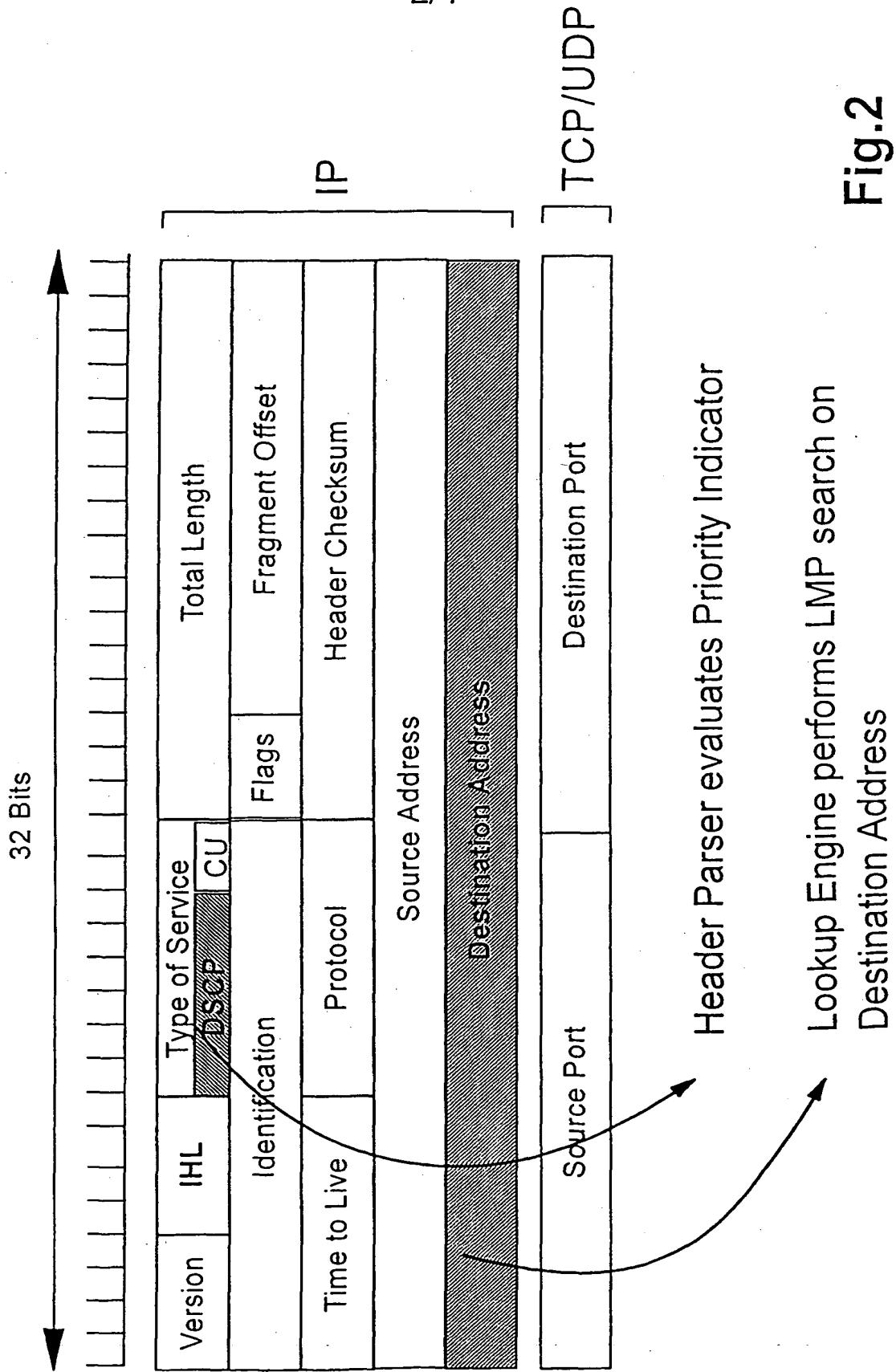
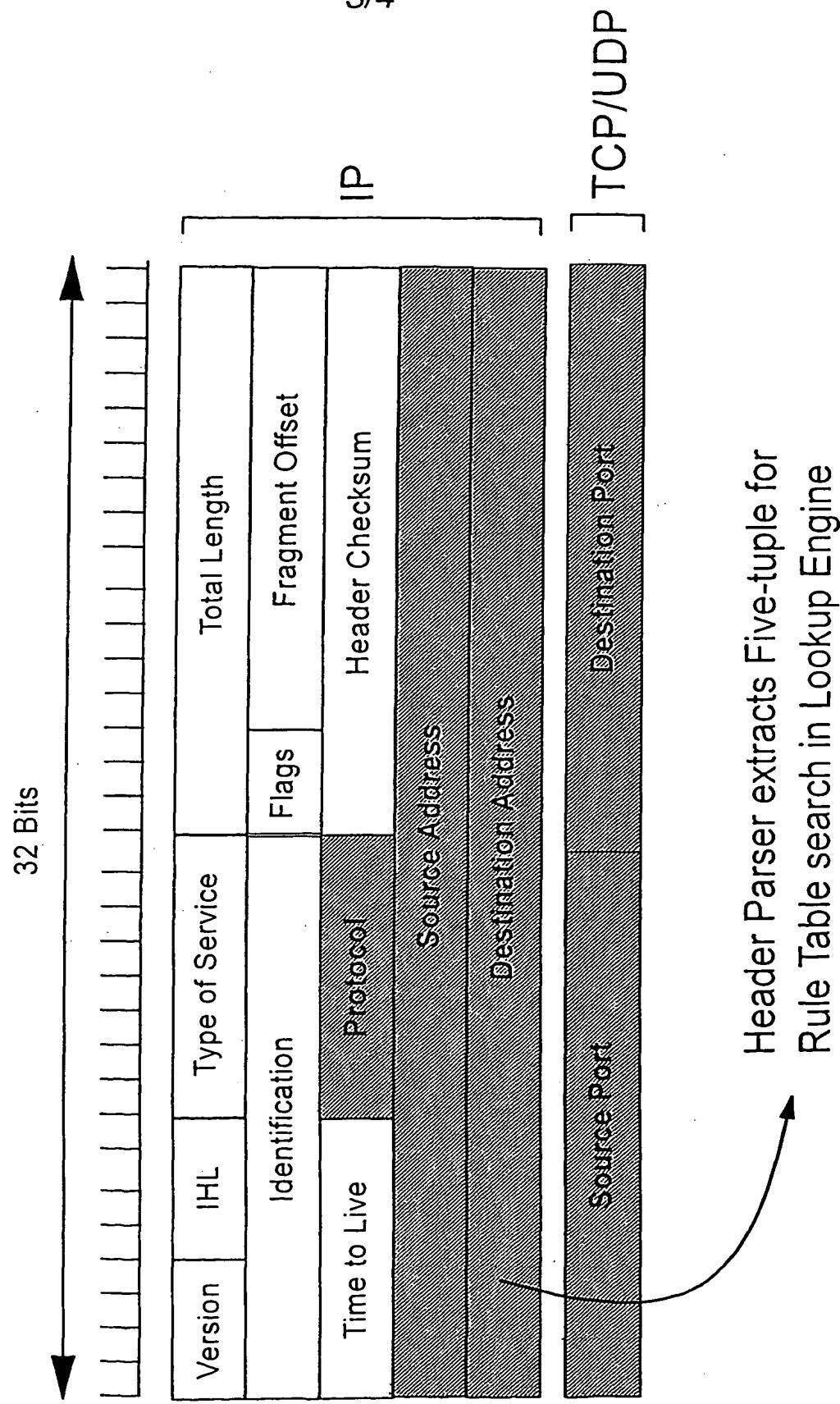


Fig.2

3/4

Fig.3



Header Parser extracts Five-tuple for
Rule Table search in Lookup Engine

4/4

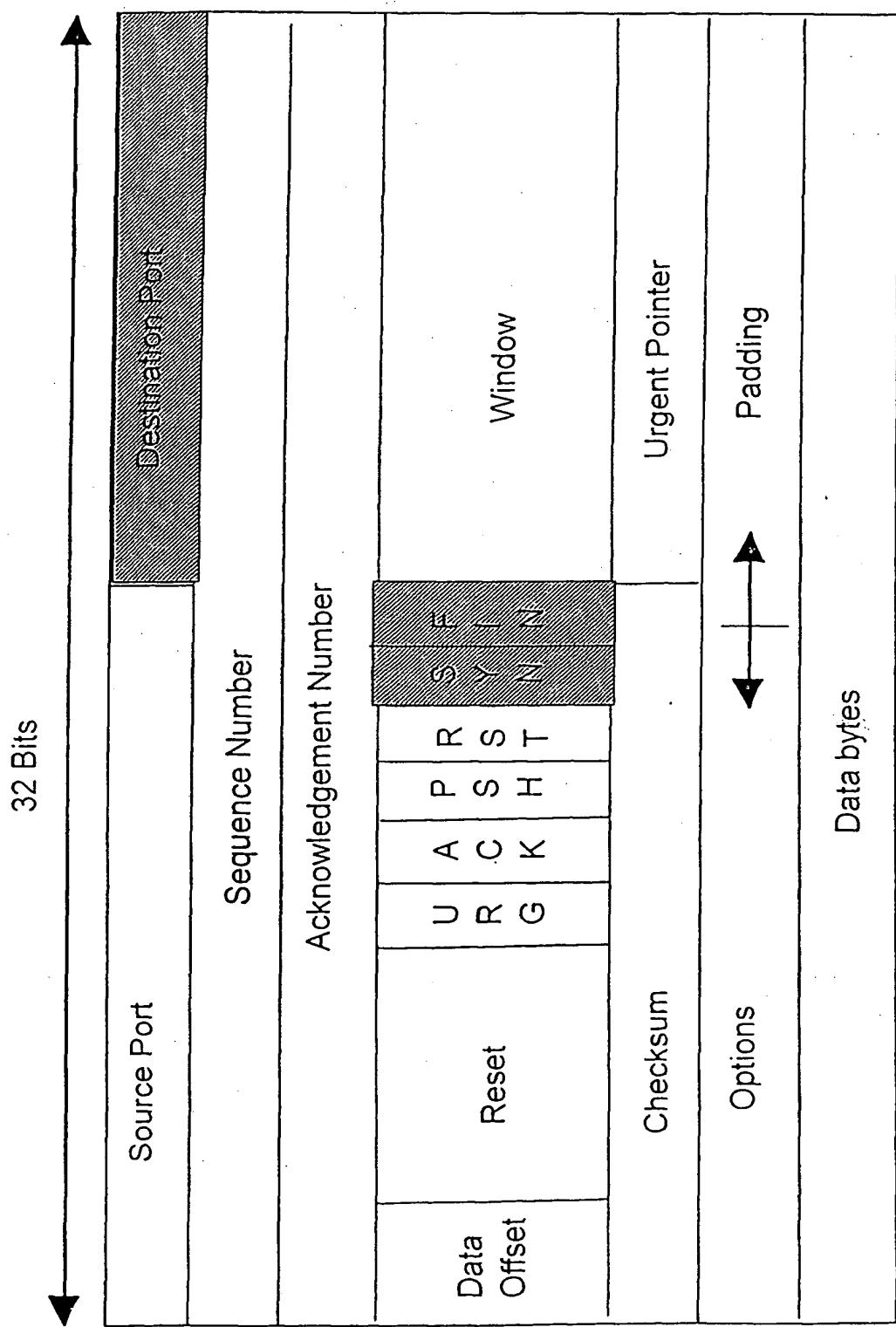


Fig.4

TCP Message with Cache Entry Selection Criteria